Specific (minor) comments

The advantage of having multiple meteorological data sources could be highlighted better. There are more advantages for sure (e.g. the possibility to fill gaps in the hourly radar data with the interpolated station data), but especially deep learning methods can make use of multiple forcing data sources at once (see Kratzert et al., 2021), which could be especially useful when mixing radar and station data as these data sources have different strengths.

- L29-L30: The DOI URL is not working (Error: DOI not found), maybe this needs to be updated? → This is a formatting error, when I click on the link, the URL ends before the line break (the DOI URL in the Data availability section works)
- L52: You could add that the meteorological timeseries are spatially aggregated to the catchment boundaries, I think this would be more clear then “catchment timeseries”.
- L84: Is this the correct citation of the exactextract package? Also in the References section (L514). I think extactextract was not published yet, so you would have to refer to it like a website? Maybe refer to one of the Github releases or to a version on pypi?
- L85: suggestion: “used to extract and spatially aggregate"
- L86: I would change “by” to “while” here
- L154: Label c) is missing in Figure 2
- L168: I think it would be good to add information about the timezone of your data. Maybe this is not strictly necessary for Great Britain as it is in GMT / UTC+0, but once there are more CAMELS datasets in sub-daily resolution, the time zone information becomes very important for interoperability. So I would also add that information for CAMELS-GB, just to be clear about it. I guess that you also do not change between summer and winter time, that information should also be added.
- L258: “Figure 4” is bold
- Figure 4: The text size of the legend is very small and should be increased.
- L276-L279: The length of the hydro-meteorological data was also extended in v2, so the hydrological and meteorological attributes were also calculated for that longer period, right? This could be added here.
  A sentence about why these attributes were calculated for the daily data only and not the hourly data could also be added (and what this implies / do you expect major differences when calculating these attributes based on hourly or daily data?). Could also be added to section 5.2 / 5.3.
- L324: The decrease in urban land cover from 2021 to 2022 in Figure S9 is really suspicious. Do you think this is an error in the data? Maybe you could reach out to the dataset’s authors for clarification and if you get a timely response you could add the explanation to the manuscript.
- L358-L372: I think the comparison of the maximum gauged flow and the maximum flows in the timeseries is very important and can give a good estimate of the uncertainties in high flows. But maybe you could add a sentence about why this is important for uncertainty estimation somewhere at the start of the section. It is about the rating curve and at which point extrapolating beyond the measured values of the rating curve begins, right?

- Table 1: “No longer providing wind speed, humidity, short-wave radiation and long-wave radiation as these were rarely used.” → it’s okay if you made this decision, but in my opinion this information could be very valuable in certain tasks, so I do not really understand why you do not provide this information anymore, if it is still available.
- S4: Labels a), b), c) are missing

Technical corrections

- L24: “1970- 2022” → missing white space, should be “1970 - 2022”
- L45: “finable” → “findable”
- L61: the second “are” is incorrect here
- L120 and L127: time spans are not consistent, e.g. “1890 – 2019” and “1961-2019”, also the case in other Lines
- L263: You suddenly use “CAMELS-GB-v2” in this chapter, to be consistent this should be changed to “CAMELS-GB v2” (also in the caption of Figure 4)
- L264: “aggregated to monthly” → “aggregated to monthly values”
- L300: “low date” should be “flow date”
- L289, L327, L340, L355, L357: citation format (remove “,” before year) “Coxon et al., (2020)” → “ Coxon et al. (2020)”, also in Table 5: “Harrigan et al., (2018)” → “Harrigan et al. (2018)” and “Salwey et al., (2023)” → “Salwey et al. (2023)”
- L392: “is based the same” → “is based on the same”
- L447: “dataset available” → “dataset is available”

Dataset

I opened all of the CSV files in CAMELS-GB v2 in Python with pandas and the catchment boundary shapefile with geopandas and did some basic checks (e.g. all IDs are present everywhere). I also visually inspected all attribute CSV files and some of the timeseries CSV files. During these checks I did not encounter any issues, the data is very nicely formatted and, together with the data description and the paper, very easy to work with, great work!

I think that the accessibility of the dataset could be enhanced. When I click on “Download the data”, I end up at the eidchub datastore where I can download individual files manually by clicking on them. This is only helpful if users are only interested in individual stations or attributes, which would not be very common. Usually, users of CAMELS-GB want to download the entire dataset. To achieve this, the user has to use wget (bulk download), which can already be challenging for users who are not very familiar with programming and/or Linux and HTTPS servers. I had to do some research on Google to get all the correct wget options to download the entire datasets and exclude the index.html files and I would consider myself quite experienced with things like this. If interested users cannot manage to easily download the dataset, they will just turn away. So I think having the option to just download the entire CAMELS-GB v2 dataset e.g. as a zip file would be very helpful and a lot more user friendly. The compressed zip file of CAMELS-GB v2 is also only 1.8 GB.
Maybe I missed something and the option to easily download the entire dataset in one go already exists, in this case it should be directly visible on the data centre page.